Second Semester Examination
2017/2018 Academic Session

May / June 2018

MST565 - Linear Models
(Model Linear)

Duration : 3 hours
[Masa : 3 jam]

Please check that this examination paper consists of **ELEVEN (11)** pages of printed material before you begin the examination.

*[Sila pastikan bahawa kertas peperiksaan ini mengandungi **SEBELAS (11)** muka surat yang bercetak sebelum anda memulakan peperiksaan ini.]*

Instructions : Answer **all four (4)** questions.

*[**Arahan :** Jawab **semua empat (4)** soalan.]*

In the event of any discrepancies, the English version shall be used.

[Sekiranya terdapat sebarang percanggahan pada soalan peperiksaan, versi Bahasa Inggeris hendaklah diguna pakai].

...2/-

Question 1

(a) Let $\mathbf{X}' = (X_1, X_2, X_3)$ be a random vector which is distributed as $N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where

$$\boldsymbol{\mu} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \quad , \boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 3 \end{pmatrix}.$$

(i) Find the marginal distribution of $\mathbf{Y}' = (X_1, X_2)$.

(ii) Find the conditional distribution of $(X_1, X_2 \mid X_3)$.

(iii) Find $\sigma_{11\square 3}$, $\sigma_{12.3}$.

(iv) Find $\rho_{12}, \rho_{31}, \rho_{12\square 3}$.

(v) Find the covariance of Z_1 and Z_2, where

$$Z_1 = 2X_1 - X_2 + X_3$$
$$Z_2 = X_1 + X_2 - X_3$$

[25 Marks]

(b) Table 1.1 shows the Ascorbic Acid (mg/100g) in one fruit juice for three packaging methods (A, B and C):

Table 1.1: Ascorbic Acid (mg/100g) for Three Packaging Methods

	A	B	C
	14.29	20.06	20.04
	19.10	20.64	26.23
	19.09	18.00	22.74
	16.25	19.56	24.04
	15.09	19.47	23.37
	16.61	19.07	25.02
	19.63	18.38	23.27
Total	120.06	135.18	164.71
Mean	17.15	19.31	23.53

Consider the analysis of variance model to represent the relationship between ascorbic acid and the three packaging methods. Write and fit the ANOVA model to test for the equality means of the three packaging methods. State your conclusion. [8 marks]

...3/-

Soalan 1

(a) *Biar* $\mathbf{X}' = (X_1, X_2, X_3)$ *sebagai suatu vektor rawak yang bertaburan* $N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$*, yang mana*

$$\boldsymbol{\mu} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \quad , \boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 3 \end{pmatrix}.$$

(i) *Cari taburan sut bagi* $\mathbf{Y}' = (X_1, X_2)$.

(ii) *Cari taburan bersyarat bagi* $(X_1, X_2 \mid X_3)$.

(iii) *Cari* $\sigma_{11\square 3}$, $\sigma_{12.3}$.

(iv) *Cari* $\rho_{12}, \rho_{31}, \rho_{12\square 3}$.

(v) *Dapatkan kovarians bagi* Z_1 *dan* Z_2*, dengan*

$$Z_1 = 2X_1 - X_2 + X_3$$
$$Z_2 = X_1 + X_2 - X_3$$

[25 markah]

(b) *Data dalam Jadual 1.1 menunjukkan kandungan Asid Askorbik (mg/100g) dalam jus buah bagi tiga kaedah pembungkusan (A, B dan C).*

Jadual 1.1: Asid Askorbik (mg/100g) bagi Tiga Kaedah Pembungkusan

	A	***B***	***C***
	14.29	*20.06*	*20.04*
	19.10	*20.64*	*26.23*
	19.09	*18.00*	*22.74*
	16.25	*19.56*	*24.04*
	15.09	*19.47*	*23.37*
	16.61	*19.07*	*25.02*
	19.63	*18.38*	*23.27*
Total	*120.06*	*135.18*	*164.71*
Mean	*17.15*	*19.31*	*23.53*

Pertimbangkan model analisis varian untuk mewakili hubungan antara asid askorbik dengan tiga kaedah pembungkusan tersebut. Tulis dan suaikan model ANOVA bagi menguji kesamaan min bagi ketiga-tiga kaedah pembungkusan itu. Nyatakan kesimpulan anda. *[8 markah]*

Question 2

(a) If $\mathbf{y}$ is $N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where

$$\boldsymbol{\mu} = \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix} \quad , \quad \boldsymbol{\Sigma} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Find a symmetric matrix $\mathbf{A}$ such that $\mathbf{y}'\mathbf{A}\mathbf{y}$ is $\chi^2(3, \lambda)$. Then find the value of noncentrality parameter λ .

[7 marks]

(b) The data in Table 2.1 relate grams plant dry weight y to percent soil organic matter x_1 and kilograms of supplemental soil nitrogen added per 1,000 square metres x_2 .

Table 2.1. Plant Dry Weight Data

y	x_1	x_2
78.5	7	2.6
74.3	1	2.9
104.3	11	5.6
87.6	11	3.1
95.9	7	5.2
109.2	11	5.5
102.7	3	7.1

(i) Propose a suitable model for the data.
(ii) What is the $\mathbf{X}$ matrix for such a model?
(iii) Find $\mathbf{X}'\mathbf{X}$, $(\mathbf{X}'\mathbf{X})^{-1}$ and $\mathbf{X}'\mathbf{y}$.
(iv) Find the least squares estimator for $\boldsymbol{\beta}$.
(v) Write the fitted regression equation.

[20 marks]

Soalan 2

(a) *Jika* $\mathbf{y}$ *adalah* $N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$*, dengan*

$$\boldsymbol{\mu} = \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix} \quad , \quad \boldsymbol{\Sigma} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix},$$

Cari matrik $\mathbf{A}$ *yang simetri supaya* $\mathbf{y}'\mathbf{A}\mathbf{y}$ *adalah* $\chi^2(3, \lambda)$*. Kemudian cari nilai parameter tidak memusat,* λ*.*

[7 marks]

(b) *Data dalam Jadual 2.1 menghuraikan* y *gram berat kering tanaman terhadap* x_1 *peratus bahan organik tanah dan* x_2 *kilogram tanah tambahan bernitrogen per seribu meter persegi.*

Jadual 2.1. Data Berat Kering Tumbuhan

y	x_1	x_2
78.5	*7*	*2.6*
74.3	*1*	*2.9*
104.3	*11*	*5.6*
87.6	*11*	*3.1*
95.9	*7*	*5.2*
109.2	*11*	*5.5*
102.7	*3*	*7.1*

(i) *Cadangkan suatu model yang sesuai untuk data tersebut.*

(ii) *Apakah matrik* $\mathbf{X}$ *bagi model tersebut?*

(iii) *Cari* $\mathbf{X}'\mathbf{X}$ *,* $(\mathbf{X}'\mathbf{X})^{-1}$ *dan* $\mathbf{X}'\mathbf{y}$*.*

(iv) *Cari penganggar kuasa dua terkecil bagi* $\boldsymbol{\beta}$*.*

(v) *Tulis persamaan regresi tersuai tersebut.*

[20 markah]

Question 3

(a) In an effort to obtain maximum yield in a chemical reaction, the value of the following variables were chosen by the experimenter:

x_1 = temperature (^0C)

x_2 = concentration of a reagent (%)

x_3 = time of reaction (hours)

with one response variable was observed:

y = percent of converted to the desired product.

The data are in the following Table 3.1:

Table 3.1. Chemical Reaction Data

y	x_1	x_2	x_3
41.5	162	23	3
33.8	162	23	8
27.7	162	30	5
21.7	162	30	8
19.9	172	25	5
15.0	172	25	8
12.2	172	30	5
4.3	172	30	8
19.3	167	27.5	6.5
6.4	177	27.5	6.5
37.6	157	27.5	6.5
18.0	167	32.5	6.5
26.3	167	22.5	6.5
9.9	167	27.5	9.5
25.0	167	27.5	3.5
14.1	177	20	6.5
15.2	177	20	6.5
15.9	160	34	7.5
19.6	160	34	7.5

Consider model $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$.

(i) Find and estimate $\text{cov}\left(\hat{\boldsymbol{\beta}}\right)$.

...7/-

(ii) Find R^2 and R_a^2 (adjusted R^2) for the model.

(iii) Test the hypothesis $H_0 : 2\beta_1 = 2\beta_2 = \beta_3$ by expressing the H_0 in the form $H_0 : \mathbf{C\beta} = \mathbf{0}$ versus $H_1 : \mathbf{C\beta} \neq \mathbf{0}$.

[15 marks]

(b) Using the same data set in Table 3.1, now consider model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_2^2 + \varepsilon .$$

Compare the fit of the model (full model) $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$ to the reduced model $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1^* + \boldsymbol{\varepsilon}^*$ from part (a) above.

[7 marks]

Soalan 3

(a) Dalam usaha untuk mendapatkan hasil maksimum dalam suatu tindakbalas kimia, nilai pemboleh ubah berikut telah dipilih oleh seorang pengkaji:

x_1 *= suhu (*^0C*)*

x_2 *= kepekatan reagen (%)*

x_3 *= masa tindakbalas (jam)*

dengan satu pemboleh ubah respon diperhatikan:

y *= peratusan perubahan bahan ke produk yang dikehendaki*

Data adalah dalam ***Jadual 3.1*** *yang berikut:*

Jadual 3.1. Data Tindakbalas Kimia

y	x_1	x_2	x_3
41.5	162	23	3
33.8	162	23	8
27.7	162	30	5
21.7	162	30	8
19.9	172	25	5
15.0	172	25	8
12.2	172	30	5
4.3	172	30	8
19.3	167	27.5	6.5
6.4	177	27.5	6.5
37.6	157	27.5	6.5
18.0	167	32.5	6.5
26.3	167	22.5	6.5
9.9	167	27.5	9.5
25.0	167	27.5	3.5
14.1	177	20	6.5
15.2	177	20	6.5
15.9	160	34	7.5
41.5	160	34	7.5

Pertimbangkan model $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$.

(i) Cari dan anggar $\text{cov}\left(\hat{\boldsymbol{\beta}}\right)$.

(ii) *Cari* R^2 *dan* R_a^2 *(*R^2 *terlaras) untuk model itu.*

(iii) *Uji hipotesis* $H_0 : 2\beta_1 = 2\beta_2 = \beta_3$ *dalam bentuk* $H_0 : \mathbf{C\beta} = \mathbf{0}$ *lawan* $H_1 : \mathbf{C\beta} \neq \mathbf{0}$.

[15 markah]

(b) *Menggunakan set data dalam Jadual 3.1*, sekarang pertimbangkan model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_2^2 + \varepsilon .$$

Bandingkan kesuaian model tersebut (model penuh) $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$ *dengan model bentuk terturunkan* $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1^* + \boldsymbol{\varepsilon}^*$ *dari bahagian (a) di atas.*

[7 markah]

Question 4

Table 4.1 shows plasma inorganic phosphate levels (mg/dl) one hour after a standard glucose tolerance test for obese subjects, with or without hyperinsulinemia, and controls (data from Jones, 1987).

Table 4.1. Plasma Phosphate Levels in Obese and Control Subjects.

Hyperinsulinemic Obese	Non-hyperinsulinemic Obese	Controls
2.3	3.0	3.1
4.1	4.1	2.6
4.2	3.9	3.1
4.0	3.1	2.2
4.6	3.3	2.1
4.6	2.9	2.4
3.8	3.3	2.8
5.2	3.9	3.4
3.1		2.9
3.7		2.6
3.8		3.1
		3.2

(a) Assuming that the one-way classification model with fixed effects is appropriate, use these data to test hypothesis that there are no mean differences among the three groups. What is your conclusion?

[10 marks]

(b) Obtain a 95% confidence interval for the difference in means between the two obese groups and comment on the interval.

[8 marks]

Soalan 4

Jadual 4.1 menunjukkan tahap fosfat tak organik plasma (mg/dl) sejam selepas suatu ujian piawai bagi toleransi glukosa terhadap subjek-subjek yang berat badan berlebihan dengan atau tanpa hiperinsulinemik dan subjek kawalan (data daripada Jones, 1987).

Jadual 4.1. Tahap Fosfat dalam Plasma Subjek yang Berat Badan Berlebihan dan Subjek Kawalan .

Hyperinsulinemic Obes	*Tanpa-hyperinsulinemic Obes*	*Kawalan*
2.3	*3.0*	*3.1*
4.1	*4.1*	*2.6*
4.2	*3.9*	*3.1*
4.0	*3.1*	*2.2*
4.6	*3.3*	*2.1*
4.6	*2.9*	*2.4*
3.8	*3.3*	*2.8*
5.2	*3.9*	*3.4*
3.1		*2.9*
3.7		*2.6*
3.8		*3.1*
		3.2

(a) *Andaikan bahawa pengkelasan sehala dengan kesan tetap adalah sesuai, gunakan data tersebut untuk menguji hipotesis bahawa tidak wujud perbezaan min antara ketiga-tiga kumpulan. Apakah kesimpulan anda?*

[10 markah]

(b) *Dapatkan selang keyakinan 95% bagi perbezaan min antara dua kumpulan subjek yang berat berlebihan dan berikan komen tentang selang keyakinan tersebut.*

[8 markah]

-oooOOooo-